

Pelatihan Analisis Instrumen Menggunakan Quest untuk Meningkatkan Kompetensi Guru dalam Evaluasi Pembelajaran

**Awal Mulia Rejeki Tumanggor^{*1}, Liza Septia Ahmad², Cristi Ascika Sekeon³,
Yohansen Frando Hadinata Silaban⁴, Isna Pujianti⁵, Yustinus Tarigan⁶**

^{1,2,3,4,5}Pendidikan Fisika, Universitas Negeri Manado, Minahasa

⁶Pendidikan Jasmani Kesehatan dan Rekreasi, Universitas Negeri Manado, Minahasa

e-mail: ^{*1}awaltumanggor@unima.ac.id, ²lizaseptia@unima.ac.id,

³casekeon@unima.ac.id, ⁴yohansensilaban@unima.ac.id, ⁵isnapujianti@unima.ac.id,

⁶yustinustarigan@unima.ac.id

Abstrak

Kegiatan pengabdian kepada masyarakat ini bertujuan untuk meningkatkan kompetensi guru dalam menganalisis kualitas instrumen evaluasi pembelajaran melalui pelatihan penggunaan software Quest berbasis Model Rasch. Sasaran kegiatan adalah 25 guru SMA di wilayah Tondano yang dilaksanakan secara daring melalui platform Zoom Meeting selama dua hari dengan pendekatan *participatory training approach*. Metode pelaksanaan mencakup tahap persiapan, pelaksanaan (penguatan konseptual, demonstrasi, praktik langsung, dan diskusi reflektif), serta evaluasi dan pendampingan. Data dikumpulkan melalui *pretest-posttest*, angket, dan observasi, kemudian dianalisis secara deskriptif. Hasil kegiatan menunjukkan peningkatan signifikan pada kompetensi peserta, dengan rerata skor *pretest* 46,80 meningkat menjadi 81,20 pada *posttest* (peningkatan 73,50%). Perhitungan N-Gain diperoleh nilai 0,65 yang termasuk kategori sedang, mengindikasikan efektivitas pelatihan yang bermakna. Sebanyak 92% peserta mengalami peningkatan substansial dan mampu mengoperasikan Quest serta menginterpretasikan output berupa tingkat kesukaran butir (logit), statistik kesesuaian butir (INFIT/OUTFIT MNSQ), dan reliabilitas. Respons peserta sangat positif, dengan 88% menyatakan materi relevan dan 84% berkomitmen mengaplikasikan hasil pelatihan di sekolah. Kegiatan ini berkontribusi pada penguatan kompetensi pedagogik guru dalam evaluasi pembelajaran berbasis bukti.

Kata kunci: analisis instrumen, Quest, Model Rasch, kompetensi guru, evaluasi pembelajaran

Abstract

This community service activity aimed to improve teachers' competency in analyzing learning evaluation instruments through training on Quest software based on the Rasch Model. The participants were 25 high school teachers in Tondano, conducted online via Zoom over two days using a participatory training approach. The method included preparation, execution (conceptual reinforcement, demonstration, hands-on practice, and reflective discussion), and evaluation stages. Data were collected through pretests-posttests, questionnaires, and observations, then analyzed descriptively. Results showed a significant increase in competency, with the average pretest score of 46.80 increasing to 81.20 on the posttest (a 73.50% increase). The N-Gain value of 0.65 falls into the medium category, indicating meaningful training effectiveness. A total of 92% of participants showed substantial improvement and were able to operate Quest and interpret outputs such as item difficulty (logit), item fit statistics (INFIT/OUTFIT MNSQ), and reliability. Participant responses were very positive, with 88% stating the material was relevant and 84% committing to apply the training results in their schools. This activity contributes to strengthening teachers' pedagogical competence in evidence-based learning evaluation.

Keywords: instrument analysis, Quest, Rasch Model, teacher competency, learning evaluation

PENDAHULUAN

Evaluasi pembelajaran merupakan komponen esensial dalam proses pendidikan karena berfungsi sebagai dasar pengambilan keputusan pedagogis, perbaikan pembelajaran, serta penentuan capaian kompetensi peserta didik. Kualitas evaluasi sangat ditentukan oleh mutu instrumen yang digunakan. Instrumen yang tidak memenuhi kriteria validitas dan reliabilitas berpotensi menghasilkan informasi yang bias dan kurang akurat dalam proses pengambilan keputusan pendidikan (Nirmala et al., 2021; Nitko & Brookhart, 2014; Tumanggor et al., 2021).

Secara regulatif, pentingnya asesmen yang bermutu ditegaskan dalam Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 23 Tahun 2016 Tentang Standar Penilaian Pendidikan yang menyatakan bahwa penilaian hasil belajar oleh pendidik harus dilakukan secara sahih, objektif, adil, dan akuntabel. Selain itu, kebijakan asesmen dalam kerangka Kurikulum Merdeka melalui Peraturan Menteri Pendidikan Dan Kebudayaan Nomor 21 Tahun 2016 Tentang Standar Isi Pendidikan Dasar Dan Menengah menekankan bahwa asesmen harus memberikan informasi yang bermakna untuk perbaikan pembelajaran dan peningkatan kompetensi peserta didik. Regulasi tersebut menuntut guru tidak hanya melaksanakan penilaian, tetapi juga mampu melakukan analisis terhadap kualitas instrumen yang digunakan.

Dalam praktiknya, banyak guru masih melakukan analisis instrumen secara terbatas menggunakan pendekatan klasik (*Classical Test Theory/CTT*), seperti menghitung tingkat kesukaran dan daya beda secara sederhana. Pendekatan ini memiliki keterbatasan karena parameter butir sangat bergantung pada karakteristik sampel, sehingga kurang stabil dan kurang akurat untuk mengukur kemampuan peserta didik secara objektif (Hambleton et al., 1991). Keterbatasan ini berpotensi menghambat implementasi prinsip objektivitas dan akuntabilitas sebagaimana diamanatkan dalam regulasi nasional.

Salah satu pendekatan yang direkomendasikan dalam evaluasi pendidikan modern adalah Model Rasch, yang merupakan bagian dari *Item Response Theory (IRT)*. Model Rasch memungkinkan estimasi parameter kemampuan peserta didik dan tingkat kesukaran butir pada skala yang sama (logit), serta menyediakan informasi mengenai kesesuaian butir (*fit statistics*) dan reliabilitas model secara komprehensif (Bond et al., 2021; Wright & Stone, 1979). Keunggulan utama Model Rasch terletak pada sifat pengukuran yang lebih objektif dan relatif invariant terhadap sampel maupun perangkat tes, sehingga mendukung praktik asesmen yang sahih dan akuntabel.

Meskipun demikian, literasi guru terhadap analisis instrumen berbasis Model Rasch masih relatif terbatas. Banyak guru belum familiar dengan perangkat lunak analisis seperti Quest, serta belum memahami interpretasi output seperti INFIT, OUTFIT, dan *Wright map*. Padahal, dalam konteks profesionalisme guru, kemampuan melakukan analisis instrumen merupakan bagian dari kompetensi pedagogik dan praktik pengambilan keputusan berbasis data (*data-driven decision making*) (Mandinach & Gummer, 2016).

Program Quest merupakan salah satu perangkat lunak yang dirancang untuk analisis data berbasis Model Rasch secara aplikatif dalam konteks pendidikan. Melalui pelatihan yang sistematis dan berbasis praktik langsung,

guru dapat memahami prosedur analisis, menginterpretasikan output, serta mengidentifikasi butir soal yang perlu direvisi atau dieliminasi (Lumembang et al., 2025; Tarigan et al., 2025). Dengan demikian, pelatihan ini tidak hanya mendukung peningkatan kompetensi teknis, tetapi juga selaras dengan tuntutan regulasi nasional mengenai penilaian yang objektif dan bermakna. Berdasarkan uraian tersebut, diperlukan kegiatan pengabdian kepada masyarakat dalam bentuk pelatihan analisis instrumen menggunakan Quest untuk meningkatkan kompetensi guru dalam evaluasi pembelajaran. Kegiatan ini diharapkan mampu memperkuat kapasitas guru dalam melaksanakan asesmen yang sahih, objektif, dan berbasis data sebagaimana diamanatkan dalam kebijakan pendidikan nasional. Adapun tujuan kegiatan ini adalah meningkatkan kompetensi guru dalam melakukan analisis instrumen evaluasi pembelajaran menggunakan program Quest berbasis Model Rasch.

METODE

Kegiatan pengabdian kepada masyarakat ini dilaksanakan dalam bentuk pelatihan berbasis *workshop* secara daring yang bertujuan untuk meningkatkan kompetensi guru dalam melakukan analisis instrumen evaluasi pembelajaran menggunakan program Quest berbasis Model Rasch. Pendekatan yang digunakan adalah *participatory training approach*, yang mengintegrasikan pemaparan konsep, demonstrasi penggunaan perangkat lunak, praktik langsung, serta diskusi interaktif (Chambers, 2014). Pendekatan ini dipilih agar peserta tidak hanya memahami aspek teoretis, tetapi juga mampu mengaplikasikan analisis instrumen secara mandiri.

Sasaran kegiatan adalah 25 guru SMA di wilayah Tondano yang mengikuti pelatihan secara daring melalui platform *Zoom Meeting*. Kegiatan dilaksanakan selama dua hari dengan total durasi delapan jam pelatihan yang mencakup sesi penguatan konseptual dan praktik analisis data. Sebelum pelaksanaan pelatihan, tim pengabdian melakukan analisis kebutuhan melalui survei awal untuk mengidentifikasi tingkat pemahaman guru terkait analisis instrumen dan pengalaman mereka dalam menggunakan perangkat lunak analisis data. Hasil survei tersebut menjadi dasar dalam penyusunan modul pelatihan yang mencakup konsep dasar evaluasi pembelajaran, pengantar Model Rasch, prinsip analisis instrumen, serta panduan operasional penggunaan program Quest. Selain itu, disusun pula instrumen evaluasi kegiatan berupa soal *pretest* dan *posttest*, angket respons peserta, serta lembar observasi partisipasi.

Pelaksanaan pelatihan diawali dengan sesi penguatan konseptual yang membahas prinsip validitas dan reliabilitas instrumen, keterbatasan pendekatan *Classical Test Theory*, serta pengenalan konsep dasar Model Rasch, termasuk interpretasi parameter logit. Selanjutnya, peserta diperkenalkan pada program Quest, mulai dari struktur data yang diperlukan, prosedur instalasi, hingga langkah-langkah menjalankan analisis. Pada sesi praktik, peserta menggunakan dataset contoh untuk melakukan input data respons siswa, menjalankan analisis Rasch, serta menginterpretasikan output yang dihasilkan, seperti nilai *item difficulty*, statistik kesesuaian butir (INFIT dan OUTFIT), serta nilai reliabilitas person dan reliabilitas item. Kegiatan diakhiri dengan sesi diskusi dan refleksi untuk mengidentifikasi butir soal yang tidak sesuai (*misfit*) serta merumuskan

langkah perbaikan instrumen.

Data dalam kegiatan ini dikumpulkan melalui dua pendekatan, yaitu kuantitatif dan kualitatif. Data kuantitatif diperoleh dari skor *pretest* dan *posttest* yang digunakan untuk mengukur peningkatan kompetensi peserta sebelum dan sesudah pelatihan. Sementara itu, data kualitatif diperoleh dari angket respons peserta dan hasil observasi selama kegiatan berlangsung. Analisis data dilakukan secara deskriptif dengan membandingkan rata-rata skor *pretest* dan *posttest*, menghitung persentase peningkatan, serta mendeskripsikan perubahan pemahaman peserta terhadap analisis instrumen berbasis Model Rasch. Respons peserta dianalisis dengan menghitung persentase tingkat kepuasan dan mengelompokkan tanggapan terbuka ke dalam tema-tema tertentu.

Data kuantitatif dari skor *pretest* dan *posttest* dianalisis secara deskriptif dengan membandingkan rata-rata skor, menghitung persentase peningkatan, serta menganalisis N-Gain (Hake, 1999) untuk mengetahui efektivitas pelatihan dengan rumus:

$$N - Gain = \frac{\text{Skor Posttest} - \text{Skor Pretest}}{\text{Skor Maksimum} - \text{Skor Pretest}}$$

Kriteria N-Gain dikategorikan sebagai berikut: tinggi (N-Gain > 0,7), sedang ($0,3 \leq \text{N-Gain} \leq 0,7$), dan rendah (N-Gain < 0,3). Data kualitatif dari angket terbuka dan lembar observasi dianalisis dengan teknik tematik, yaitu mengelompokkan tanggapan ke dalam tema-tema tertentu seperti kemudahan penggunaan software, kejelasan materi, kendala teknis, atau rencana implementasi. Hasil analisis ini digunakan untuk mendeskripsikan perubahan pemahaman peserta serta mengevaluasi jalannya kegiatan secara menyeluruh.

HASIL DAN PEMBAHASAN

Pelaksanaan pelatihan analisis instrumen menggunakan program Quest diikuti oleh 25 guru SMA di wilayah Tondano yang berpartisipasi secara aktif selama kegiatan berlangsung. Evaluasi peningkatan kompetensi dilakukan melalui perbandingan skor *pretest* dan *posttest* yang dirancang untuk mengukur pemahaman konseptual mengenai evaluasi pembelajaran, prinsip dasar Model Rasch, serta kemampuan interpretasi output program Quest. Hasil analisis menunjukkan adanya peningkatan yang signifikan pada capaian peserta.

Tabel 1 Rekapitulasi Skor *Pretest* dan *Posttest* Peserta

Komponen Penilaian	Rerata <i>Pretest</i>	Rerata <i>Posttest</i>	Peningkatan	Persentase Peningkatan
Pemahaman Konsep IRT	45,20	80,50	35,30	78,10%
Operasional Quest	44,80	82,40	37,60	83,93%
Interpretasi Output	50,40	80,70	30,30	60,12%
Rata-rata Total	46,80	81,20	34,40	73,50%

Berdasarkan Tabel 1, rata-rata skor *pretest* sebesar 46,80 meningkat menjadi 81,20 pada *posttest*, dengan selisih peningkatan sebesar 34,40 poin atau sekitar 73,50% dari skor awal. Sebanyak 23 peserta (92%) menunjukkan peningkatan skor yang substansial, sementara dua peserta lainnya tetap mengalami peningkatan meskipun dalam kategori moderat. Tidak ditemukan

peserta yang mengalami penurunan skor.

Untuk menguji efektivitas pelatihan secara lebih akurat, dilakukan perhitungan N-Gain. Dengan asumsi skor maksimum 100, diperoleh:

$$N - Gain = \frac{81,20 - 46,80}{100 - 46,80} = \frac{34,40}{53,20} = 0,65$$

Nilai N-Gain sebesar 0,65 termasuk dalam kategori sedang ($0,3 \leq N\text{-Gain} \leq 0,7$). Hal ini mengindikasikan bahwa pelatihan memberikan kontribusi yang bermakna terhadap peningkatan kompetensi peserta. Temuan ini sejalan dengan hasil pelatihan serupa yang dilakukan oleh Kondo Lembang et al. (2021) yang melaporkan peningkatan pemahaman guru tentang analisis data setelah mengikuti workshop berbasis praktik.

Kemampuan Praktik dan Interpretasi Output Quest

Selama sesi praktik, peserta melakukan analisis data menggunakan dataset contoh yang telah disiapkan tim pengabdian. Pada tahap awal, sebagian peserta mengalami kesulitan dalam memahami konsep logit dan perbedaan pendekatan antara *Classical Test Theory* dan Model Rasch. Beberapa peserta mengaku baru pertama kali mendengar istilah "logit" dan bingung dengan makna nilai negatif pada tingkat kesukaran butir. Namun, melalui pendampingan bertahap dan diskusi interaktif, peserta mulai mampu membaca dan menginterpretasikan output yang dihasilkan oleh program Quest.

MISKONSEPSI SISWA				
Item Estimates (Thresholds) all on all (N = 113 L = 9 Probability Level= .50)				
2.0				
	X	3.4		
	XXX			
	XXX			
	XXXXXX	3.3		
1.0				
	XX	1.4	4.4	
	XXXXXXXXXXXX	5.3		
	XXXXXXXXXXXX	2.4		
	XXXXXXXXXXXX	1.3	2.3	4.3
	XXXXXXXXXX	2.2		
	XXXXXXXXXX			
	XXXXXXXXXX	6.4		
	XXXXXXXXXX			
.0				
	XXXXXXXXXX	5.2	8.4	
	XXXXX	9.4		
	XX	6.3		
	XXXXX	7.4	8.3	
		7.3		
	XXXX			
	X	3.2	6.2	
-1.0				
	X	8.2	9.3	
	XXXX			
-2.0				
		9.2		
	X	7.2		
-3.0				
Each X represents 1 students				

MISKONSEPSI SISWA							
Item Fit all on all (N = 113 L = 9 Probability Level= .50)							
INFIT							
MNSQ	.56	.63	.71	.83	1.00	1.20	1.40
1 item 1				*			
2 item 2					*		
3 item 3					*		
4 item 4						*	
5 item 5							*
6 item 6			*				
7 item 7					*		
8 item 8		*					
9 item 9		*					
=====							
MISKONSEPSI SISWA							
Case Estimates all on all (N = 113 L = 9 Probability Level= .50)							
Summary of case Estimates =====							
Mean	.26						
SD	.73						
SD (adjusted)	.61						
Reliability of estimate	.70						
Fit Statistics =====							
Infit Mean Square			Outfit Mean Square				
Mean	.97	Mean	1.14				
SD	.56	SD	1.60				
Infit t			Outfit t				
Mean	-.10	Mean	.10				
SD	1.07	SD	.95				
0 cases with zero scores							
0 cases with perfect scores							
=====							

Gambar 1 Tampilan Output Quest yang Dianalisis Peserta

Peserta dapat mengidentifikasi tingkat kesukaran butir dalam satuan logit, membedakan butir yang terlalu mudah ($\text{logit} < -2$) dan terlalu sulit ($\text{logit} > +2$), serta mengenali butir yang tidak sesuai (misfit) berdasarkan nilai INFIT dan OUTFIT Mean Square (MNSQ) dengan kriteria rentang ideal 0,77 - 1,30 (Sumintono & Widhiarso, 2015; Tumanggor & Supahar, 2021). Kemampuan ini menjadi indikator penting karena menunjukkan bahwa guru tidak hanya memahami aspek teknis penggunaan perangkat lunak, tetapi juga mampu melakukan interpretasi analitis terhadap kualitas instrumen yang digunakan.

Salah satu peserta dalam diskusi reflektif menyatakan, *"Selama ini saya hanya menghitung persentase jawaban benar siswa, lalu soal dianggap baik kalau persentasenya tidak terlalu tinggi atau terlalu rendah. Ternyata analisisnya jauh lebih dalam dengan Model Rasch, kita bisa tahu mana soal yang benar-benar fit dengan model dan mana yang bias."* Pernyataan ini merefleksikan perubahan paradigma dari sekadar administratif menuju evaluasi berbasis bukti.

Perubahan Perspektif tentang Evaluasi Pembelajaran

Diskusi reflektif pada akhir sesi praktik menunjukkan adanya perubahan perspektif peserta terhadap evaluasi pembelajaran. Sebelum pelatihan, sebagian besar guru masih mengandalkan analisis sederhana berupa persentase ketuntasan atau perhitungan tingkat kesukaran berdasarkan pendekatan klasik. Setelah mengikuti pelatihan, peserta menyadari bahwa Model Rasch memungkinkan analisis yang lebih objektif dan akurat karena memperhitungkan interaksi antara kemampuan peserta didik dan karakteristik butir soal dalam satu skala pengukuran yang sama.

Pemahaman ini penting dalam mendukung implementasi penilaian yang sahih, objektif, dan akuntabel sebagaimana diamanatkan dalam regulasi nasional mengenai standar penilaian pendidikan. Sebagaimana dikemukakan oleh Marlina et al. (2024), penggunaan IRT dalam analisis instrumen memberikan keunggulan dalam hal invariansi parameter, sehingga hasil analisis dapat digeneralisasi pada populasi yang lebih luas. Guru yang memahami prinsip ini akan lebih berhati-hati dalam menyusun dan memilih instrumen yang benar-benar berkualitas.

Respons Peserta terhadap Pelaksanaan Pelatihan

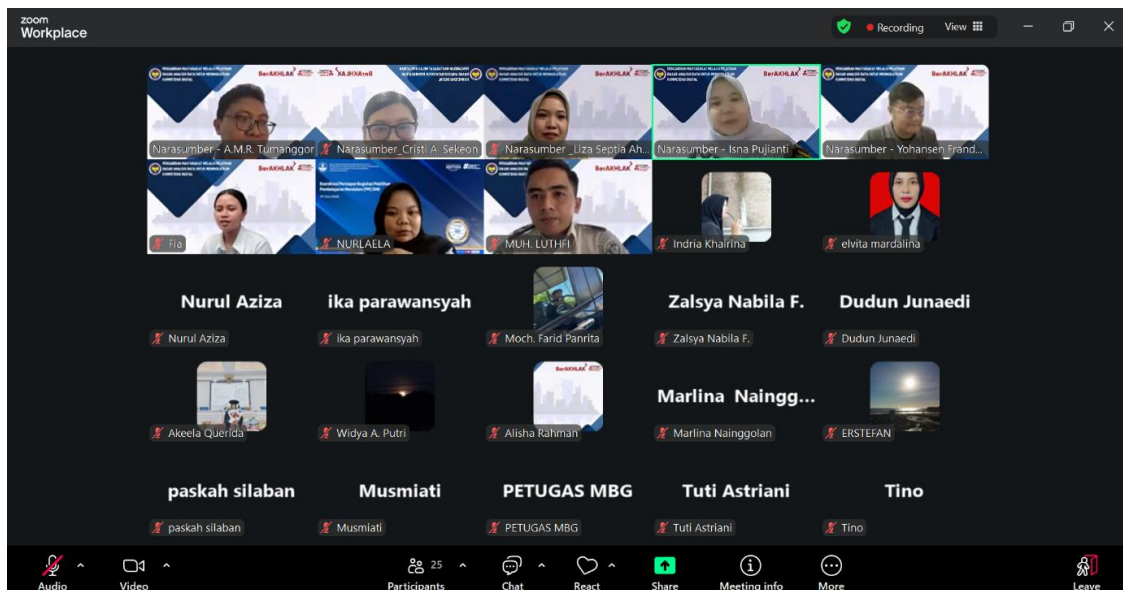
Respons peserta terhadap pelatihan juga menunjukkan hasil yang positif. Berdasarkan angket evaluasi kegiatan yang diisi oleh 25 peserta, diperoleh data sebagai berikut:

Tabel 2 Respons Peserta terhadap Pelaksanaan Pelatihan

Pernyataan	Setuju (%)	Tidak Setuju (%)
Materi pelatihan relevan dengan kebutuhan profesional	88%	12%
Pelatihan membantu memahami kualitas instrumen secara lebih komprehensif	92%	8%
Fasilitator menyampaikan materi dengan jelas dan komunikatif	90%	10%
Modul dan panduan praktis memudahkan pembelajaran	86%	14%
Berkomitmen menerapkan analisis Quest pada instrumen di sekolah	84%	16%
Pelatihan ini membuka wawasan baru tentang evaluasi berbasis data	94%	6%

Data pada Tabel 2 menunjukkan bahwa mayoritas peserta memberikan respons positif terhadap pelaksanaan pelatihan. Sebanyak 88% peserta menyatakan bahwa materi yang diberikan sangat relevan dengan kebutuhan profesional mereka, khususnya dalam penyusunan dan evaluasi soal ujian. Sebanyak 92% peserta menyatakan bahwa pelatihan membantu mereka memahami kualitas instrumen secara lebih komprehensif dibandingkan pendekatan yang selama ini digunakan. Selain itu, 84% peserta menyatakan komitmen untuk mencoba menerapkan analisis Quest pada instrumen yang digunakan di sekolah masing-masing.

Tanggapan terbuka dari peserta mengindikasikan beberapa tema utama. Pertama, pelatihan ini membuka wawasan baru tentang pentingnya evaluasi berbasis data (*data-driven assessment*). Kedua, peserta merasa mendapatkan keterampilan praktis yang dapat langsung diimplementasikan. Ketiga, beberapa peserta menyarankan perlunya pendampingan lanjutan karena masih merasa perlu bimbingan saat mencoba mandiri. Salah satu peserta menulis, "*Pelatihannya sangat bermanfaat, tapi saya khawatir lupa kalau tidak segera dipraktikkan. Semoga ada grup diskusi lanjutan.*"



Gambar 2 Dokumentasi Kegiatan Pelatihan Analisis Instrumen Tes

Refleksi Keberhasilan Pendekatan Partisipatif

Secara keseluruhan, hasil kegiatan menunjukkan bahwa pendekatan *participatory training approach* yang diterapkan efektif dalam meningkatkan kompetensi guru. Keterlibatan aktif peserta melalui praktik langsung, diskusi kelompok, serta refleksi bersama memungkinkan terjadinya proses konstruksi pengetahuan secara lebih mendalam. Pelatihan tidak hanya meningkatkan pengetahuan konseptual, tetapi juga membangun keterampilan teknis dan kepercayaan diri guru dalam melakukan analisis instrumen secara mandiri. Keberhasilan ini tidak terlepas dari desain pelatihan yang memperhatikan prinsip andragogi, di mana peserta dewasa belajar paling efektif ketika materi dikaitkan dengan pengalaman dan kebutuhan nyata mereka (Knapke et al., 2024). Dengan menggunakan dataset contoh yang relevan

dengan konteks pembelajaran di sekolah, peserta dapat melihat langsung manfaat analisis Quest dalam mengevaluasi soal-soal yang mereka hadapi sehari-hari.

Keterbatasan Kegiatan

Meskipun hasil kegiatan menunjukkan peningkatan yang signifikan, terdapat beberapa keterbatasan yang perlu dicatat. Pertama, durasi pelatihan yang relatif singkat (8 jam) belum memungkinkan pendalaman konsep IRT secara komprehensif, terutama bagi peserta yang sama sekali baru mengenal pendekatan ini. Kedua, keterbatasan interaksi langsung dalam mode daring memengaruhi dinamika diskusi kelompok dan pendampingan individual. Beberapa peserta dengan koneksi internet tidak stabil mengalami kesulitan mengikuti sesi praktik secara optimal. Ketiga, evaluasi jangka panjang belum dilakukan untuk memastikan apakah keterampilan yang diperoleh benar-benar diaplikasikan dalam praktik nyata di sekolah. Dengan demikian, kegiatan pengabdian ini berkontribusi terhadap penguatan kompetensi pedagogik guru, khususnya dalam aspek evaluasi pembelajaran berbasis bukti dan berorientasi pada peningkatan mutu instrumen.

KESIMPULAN

Kegiatan pengabdian masyarakat berupa pelatihan analisis instrumen menggunakan software Quest bagi guru SMA di wilayah Tondano telah berhasil dilaksanakan secara daring dan mencapai tujuan yang ditetapkan. Terjadi peningkatan kompetensi guru yang signifikan dalam aspek pemahaman konsep analisis instrumen berbasis Model Rasch, keterampilan mengoperasikan Quest, dan kemampuan menginterpretasi output untuk perbaikan kualitas soal. Rata-rata skor *pretest* 46,80 meningkat menjadi 81,20 pada *posttest* (peningkatan 73,50%), dengan nilai N-Gain 0,65 yang termasuk kategori sedang. Sebanyak 92% peserta menunjukkan peningkatan substansial, dan peserta mampu mengidentifikasi tingkat kesukaran butir, statistik kesesuaian butir, serta nilai reliabilitas dari output Quest. Kegiatan ini menjawab permasalahan peserta terkait keterbatasan pengetahuan dan keterampilan dalam menganalisis kualitas instrumen evaluasi menggunakan pendekatan modern. Penguasaan software Quest memungkinkan guru menghasilkan instrumen penilaian yang lebih valid, reliabel, dan akurat dalam mengukur kemampuan siswa. Respons peserta sangat positif, dengan 92% menyatakan pelatihan membantu pemahaman komprehensif tentang kualitas instrumen dan 84% berkomitmen mengaplikasikan keterampilan yang diperoleh.

DAFTAR PUSTAKA

- Bond, T. G., Yan, Z., & Heene, M. (2021). *Applying the Rasch Model, Fundamental Measurement in the Human Sciences*. Routledge. <https://doi.org/10.4324/9781410614575>
- Chambers, R. (2014). PRA, PLA and Pluralism: Practice and Theory. *The SAGE Handbook of Action Research*, 297–318. <https://doi.org/10.4135/9781848607934.n28>
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of*

- Item Response Theory Library*. Sage Publication, Inc.
- Knapke, J. M., Hildreth, L., Molano, J. R., Schuckman, S. M., Blackard, J. T., Johnstone, M., Kopras, E. J., Lamkin, M. K., Lee, R. C., Kues, J. R., & Mendell, A. (2024). Andragogy in Practice: Applying a Theoretical Framework to Team Science Training in Biomedical Research. *British Journal of Biomedical Science*, 81(March), 1–8. <https://doi.org/10.3389/bjbs.2024.12651>
- Lembang, F. K., Djami, R. J., & Latuihamallo, C. S. (2021). Peningkatan Kompetensi Guru SD Kristen 1 Hunuth Melalui Pelatihan Teknik Analisis Data Statistik Penelitian Tindakan Kelas Menjadi Karya Ilmiah. *BAKTI (Jurnal Pengabdian Kepada Masyarakat)*, 1(1), 12–16. <https://doi.org/10.51135/baktivol1iss1pp12-16>
- Lumembang, M. M., Arundaa, R., & Tumanggor, A. M. R. (2025). Optimalisasi Pembelajaran Konsep Kemagnetan melalui Pelatihan Sensor Magnetik IoT di SMPN 3 Kawangkoan. *JURNAL ABDIMAS MADUMA*, 4(3), 188–195. <https://doi.org/10.52622/jam.v4i3.554>
- Mandinach, E. B., & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366–376. <https://doi.org/10.1016/j.tate.2016.07.011>
- Marlina, M., Rahim, A., Aziz, F., & Arsalaan, A. T. (2024). Pelatihan analisis kualitas instrumen penilaian hasil belajar dengan pendekatan Classical Theory dan Item Response Theory di sekolah dasar. *KACANEGARA Jurnal Pengabdian Pada Masyarakat*, 7(2), 209. <https://doi.org/10.28989/kacanegara.v7i2.1888>
- Nirmala, M. F. T., Tumanggor, A. M. R., & Supahar. (2021). Analysis of Validity and Reliability of Diagnostic Test of Picture Representation Ability in High School Physics Learning. *Proceedings of the 7th International Conference on Research, Implementation, and Education of Mathematics and Sciences (ICRIEMS 2020)*, 528(Icriems 2020), 538–546. <https://doi.org/10.2991/assehr.k.210305.079>
- Nitko, A. J., & Brookhart, S. M. (2014). Educational Assessment of Students Sixth Edition. In *Pearson New International Edition*.
- Peraturan Menteri Pendidikan Dan Kebudayaan Nomor 21 Tahun 2016 Tentang Standar Isi Pendidikan Dasar Dan Menengah (2016).
- Permendikbud No. 23 Tahun 2016 Tentang Standar Penilaian Pendidikan (2016). <https://doi.org/10.31227/osf.io/munp2>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Trim Komunika Publishing House.
- Tarigan, Y., Tumanggor, A. M. R., Paembonan, T. L., Varisky Abraham Dumanaw, & Jein Jeilen Tampi. (2025). Sosialisasi Penggunaan SPSS dalam Meningkatkan Kompetensi Analisis Data Statistik Mahasiswa FMIPAK UNIMA. *Journal Inovasi Pengabdian Masyarakat*, 2(2), 132–141. <https://doi.org/10.65255/jipmas.v2i2.198>
- Tumanggor, A. M. R., & Supahar. (2021). Dissemination of Online Diagnostic Test Instrument to Diagnose High School Students' Mathematical Representation Ability: The Case of Work and Energy. *Proceedings of the 6th International Seminar on Science Education (ISSE 2020)*, 541, 528–537.

<https://doi.org/10.2991/assehr.k.210326.076>

- Tumanggor, A. M. R., Supahar, & Nirmala, M. F. T. (2021). The Development of Diagnostic Test Instrument for Verbal Representation Ability in High School Physics Learning. *Proceedings of the 7th International Conference on Research, Implementation, and Education of Mathematics and Sciences (ICRIEMS 2020)*, 528, 471–476.
<https://doi.org/10.2991/assehr.k.210305.069>
- Wright, B. D., & Stone, M. H. (1979). *Best Test Design: Rasch Measurement*. In Mesa Press. Routledge.